DOCUMENT RESUME

ED 310 121 TM 013 694

AUTHOR van der Linden, Wim J.; Boekkooi-Timminga, Ellen
TITLE A Zero-One Programming Approach to Gulliksen's Matched Random Subtests Method. Research Report 86-4.
INSTITUTION Twente Univ., Enschede (Netherlands). Dept. of Education.
SPONS AGENCY Netherlands Inst. for the Advancement of Pure Research, The Hague.
PUB DATE 86
NOTE 28p.; Also cited as Project Psychometric Aspects of Item Banking No. 5.
AVAILABLE FROM Mediatheek, Faculteit Toegepaste Onderwijskunde, Universiteit Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Algorithms; Computer Assisted Testing; *Computer Software; Difficulty Level; Estimation (Mathematics); Graphs; *Programing; Test Items; *Test Reliability
IDENTIFIERS Alpha Coefficient; Classical Test Theory; *Gulliksens Matched Random Subtests Method; Parallel Test Forms; *Zero One Programing

ABSTRACT

In order to estimate the classical coefficient of test reliability, parallel measurements are needed. H. Gulliksen's matched random subtests method, which is a graphical method for splitting a test into parallel test halves, has practical relevance because it maximizes the alpha coefficient as a lower bound of the classical test reliability coefficient. This paper formulates this same problem as a zero-one programming problem, the advantage being that it can be solved by algorithms already existing in computer code. Focus is on giving Gulliksen's method a sound computational basis. How the procedure can be generalized to test splits into components of any length is shown. An empirical illustration of the procedure is provided, which involves the use of the algorithm developed by A. H. Land and A. Doig (1960), as implemented in the LANDO program. Item difficulties and item-test correlations were estimated from a sample of 5,418 subjects--a sample size that is large enough to prevent capitalizing on chance in the Gulliksen method. Two data tables and one graph are provided. (Author/TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED310121

A Zero-One Programming Approach to Gulliksen's Matched Random Subtests Method

Research Report 88-4

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. NELISSEN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

W.J. van der Linden
E. Boekkooi-Timminga

TM013694

Division of Educational Measurement and Data Analysis

University of Twente

BEST COPY AVAILABLE

Department of Education

Project Psychometric Aspects of Item Banking No. 5

Colophon

Typing : A. Burchartz, L. Padberg
Cover design : M. Driessen, AV-section, University of Twente
Printed by : Central Reproduction Department, University of Twente

A Zero-One Programming Approach to Gulliksen's Matched Random Subtest Method

Wim J. van der Linden

Ellen Boekkooi-Timminga

Abstract

Gulliksen's matched random subtests method is a graphical method to split a test into parallel test halves which has practical relevance because it maximizes coefficient α as a lower bound to the classical test reliability coefficient. In the paper the same problem is formulated as a zero-one programming problem, the advantage being that it can be solved by algorithms already existing in computer code. It is shown how the procedure can be generalized to test splits into components of any length. An empirical illustration of the procedure concludes the paper.

A Zero-one Programming Approach to Gulliksen's Matched Random Subtests Method

In order to estimate the classical coefficient of test reliability, parallel measurements are needed. Methods proposed to meet this requirement in practice are retesting the same subjects with the same test after some time has elapsed or carefully constructing a parallel test and testing the same subjects with both instruments. As is known from practical experience, though, these methods do not work well. The main objection against the test-retest method is that replicate test administrations are impossible with living subjects who may exhibit all kinds of interfering processes as remembering earlier administrations, learning and forgetting between administrations, or taking a dislike to another administration. The parallel-forms method, in fact, constitutes a dilemma. It assumes that it is possible to construct two <u>different</u> tests with exactly the <u>same</u> measurement properties. Practical experience shows that this ideal may be attained to some extent but is never realized exactly.

As a possible way out of this fundamental problem, Kuder and Richardson (1937) proposed their formulas 20 and 21 which can be estimated using (dichotomous) item and test scores from a single administration. A generalization of these formulas to non-dichotomous items or test components of any lengtn is known as Cronbach's (1951) coefficient α :

$$(1) \qquad \alpha \equiv \frac{n}{n-1}\left(1 - \frac{\sum \sigma^2(Y_g)}{\sigma_X^2}\right), \qquad \sigma_X^2 > 0,\ n > 1,$$

where $\sigma^2(Y_g)$ is the variance of the scores Y_g on test component g, σ_X^2 the variance of the score X, and n the number of components. The usual choices of test components in this internal-consistency method are the individual test items or test halves. Estimates of the test reliability based on the latter are known as split-halves estimates. A generalization of (1) to any split was introduced by Raju (1977) and is known as coefficient β_K.

Analysis of the relationship of (1) to the definition of the reliability coefficient reveals that they are equal to each other only if the test components are essentially τ-equivalent ; otherwise (1) is a lower bound to the test reliability (e.g., Lord & Novick, 1968, sect. 4.4). Although this requirement is less restrictive than the one of parallel measurements, it seems to give rise to the same practical problems as for the test-retest and parallel-forms methods. However, there is a possibility of optimization that the latter methods do not possess. Since (1) is a lower bound to the reliability for <u>any</u> split of the test into components, and these bounds are not necessarily equal, we may look for the split with the greatest lower bound and base our estimation of the reliability coefficient on this. It is for this reason that the internal-consistency method has not only a practical but also some theoretical appeal.

Gulliksen (1950) proposed a method for splitting tests optimally

into halves which is know as the matched random subtests method. The method is the only one available for this important purpose and is described in most textbooks on test theory (e.g., Allen & Yen, 1979, sect. 4.4). In spite of this, it has not found its implementation in standard computer packages for test analysis and is hardly being used on a routine basis the reason being that the method is graphic and must be performed by hand. It is the purpose of this paper to present a version of Gulliksen's method that is derived from zero-one programming. Algorithms for this method exist and are amply available in computer code. In the remainder of this paper, first Gulliksen's method is outlined. Then, a zero-one programming version of the method is proposed. Next, this version is illustrated using empirical test data. The paper ends with some concluding remarks and recommendations.

Gulliksen's Matched Random Subtests Method

Gulliksen's method is usually formulated for dichotomous item scores but can easily be generalized to other situations. For dichotomous item scores, the method involves two parameters for each item, its difficulty and discriminating power. Let π_i and ρ_{iX} denote the classical definitions of these parameters. Then the former is the expected item score and the latter the point biserial correlation between the item and the test score. Each item is plotted on a graph with its values for the two parameters as coordinates. Next, pairs of items are formed, the

criterion being that each pair should have points on the graph as close to each other as possible. Test halves are obtained by assigning one randomly chosen item from each pair to one test half and the remaining items to the other.

Figure 1 shows a typical Gulliksen plot. The points are estimates

---

Insert Figure 1 about here

---

for a 20-item version of a mathematics achievement test used in the Second Mathematics Study of the International Association for the Evaluation of Education based on a Dutch sample of 5418 subjects. The same data will be used in the empirical example below. Note that some pairs in Figure 1 are obvious. Others, however, are not. Item 16, for instance, could be paired with 19 but this choice has consequences for the possibilities of 8, whereas the choice for this item, in turn, restricts the possibilities for 2, and so on. In fact, it is the absence of a clear-cut criterion for coping with such dependencies that may make the method hard to be used for larger sets of items.

Let Y_g in (1) represent the observed score on test half g which consists of n_g items (g = 1, 2). A well-known result from the classical test theory is that, for dichotomous items, the expected values and variances of Y_g can be written as functions of π_i and ρ_{iY_g} only. Assuming $\rho_{iY_g} = \rho_{iX}$ for g = 1, 2, as is

implicitly done in the Gulliksen method, the expressions are

$$(2) \qquad \mu_{Y_g} = \sum_{i=1}^{n_g} \pi_i$$

$$(3) \qquad \sigma^2_{Y_g} = [\sum_{i=1}^{n_g} \pi_i(1-\pi_i)\rho_{iX}]^2$$

Gulliksen's method is motivated by the fact that pairwise matching of the items on π_i ensures that μ_{Y_1} and μ_{Y_2} are approximately equal. Hence, a necessary condition for the two halves to have the same true scores is met. As matching on ρ_{iX} also ensures approximately equal values of (3) for g = 1, 2, the two halves may have equal error variances and meet the requirements of parallel measurements.

As already mentioned, Gulliksen's method is graphic. It supposes the presence of a judge inspecting the graph and pairing the items. It is not a algorithm in the sense that all its rules can be written in computer code. As illustrated earlier, its criterion for pairing the items is not unequivocal. Therefore, situations may arise where the judge does not know with certainty which of the possible pairs to choose. Also, the random assignment of items from pairs to test halves may be suboptimal. In particular, when the items within pairs are not close to each other there is a non-negligible probability for random assignment to result in test halves being less parallel than necessary. Another desirable improvement on the method would be an algorithm equally well applicable to splits into other components than test halves. Splits

of tests into thirds or quarters, for instance, require the division of the plots into triples or quadruples of items. It is unlikely that this can be done satisfactorily for larger tests just by inspecting plots. On the other hand, since such splits also yield values for (1) that are lower bounds to the reliability coefficient, and it seems unwise to confine the search of the greatest lower bound only to the subset of splits into test halves. As a final comment on the Gulliksen method, it is noted that, like any other method of item selection, the danger of chance capitalization may arise if it is used with sample statistics instead of parameters. For this reason, it can only be recommended as a large-sample solution to the problem of splitting a test into parallel halves. The same holds if the zero-one programming formulation of Gulliksen's method given below is used with statistics instead of parameters.

A Zero-one Programming Version of Gulliksen's Method

Gulliksen's method consists of two steps--pairing the items and assigning items from pairs to test halves. Both tasks can be performed using techniques from zero-one programming. Interest in the application of zero-one programming techniques originated recently from a paper by Theunissen (1985) who applied them to solve the problem of automated test design in item response theory. This problem is pursued further in Theunissen and Verstralen (1986) and van der Linden and Boekkooi-Timminga (1986), whereas Boekkooi-

Timminga (1986a, 1986b) provide extensions to the problems of simultaneous test design and the design of parallel tests in item response theory. The techniques used below have a close relationship to the ones in the last two references but are applied in the context of classical test theory here while also use is made of the minimax approach in van der Linden and Boekkooi-Timminga.

Pairs of items

In Gulliksen's method the items are paired on inspection. It is suggested to replace this situation by the following unequivocal criterion. In the graph the Euclidean distance

$$\delta_{ij} = [(\pi_i - \pi_j)^2 + (\rho_{iX} - \rho_{jX})^2]^{1/2} \qquad (4)$$

between the points i and j ($i \neq j$) is considered. It is proposed to pair the items such that the sum of the within-pair distances is minimal. In the following, as is necessary in the Gulliksen method, it is supposed that n is an even number. (If n is odd, one item must be deleted and a Spearman-Brown correction with factor n/(n-1) should be applied to the eventual reliability estimate). Let x_{ij} be a binary decision variable denoting whether i and j are a pair or not. That is,

$$x_{ij} = \begin{cases} 1 & \text{i and j are a pair,} \\ 0 & \text{otherwise,} \end{cases} \qquad i < j.$$

The problem is to decide on the n(n-1)/2 values of x_{ij} such that

the criterion of a minimal sum of distances is met. Now the product $\delta_{ij}x_{ij}$ is equal to the distance between i and j if they are a pair, and to zero otnerwise. The problem is thus to minimize the sum of these products subject to the constraints that each item has to be a member of exactly one pair. In the usual zero-one programming format the problem is as follows

$$(5) \qquad \text{minimize} \quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \delta_{ij}x_{ij}$$

subject to

$$(6) \qquad \sum_{i=1}^{j-1} x_{ij} + \sum_{i=j+1}^{n} x_{ji} = 1 \qquad j = 1, \ldots, n$$

$$(7) \qquad x_{ij} \in \{0, 1\} \qquad i = 1, \ldots, n-1; \quad j = i+1, \ldots, n$$

where for notational convenience the sums in (6) are equal to zero if the upper bound to the index is smaller than the lower bound, or conversely. The objective function in (5) is defined as the minimization of the sum of all within-pair distances. The constraints in (6) guarantee that for each item the decision variables x_{ij} $(i < j)$ take the value 1 exactly once, which means that each item arrives in exactly one pair. In (7) the decision variables are constrained to be binary.

The problem in (5) - (7) is a standard zero-one programming problem that is found in textbooks on linear programming (e.g.,

Wagner, 1975, chap.13). Algorithms to solve the problem can be found in the same references and have been implemented in various computer programs. In the empirical example below, the program LANDO is used which is based on the branch-and-bound method by Land and Doig (1960). The output of the program is the $n(n-1)/2$ values of the decision variables x_{ij} with $n/2$ values equal to 1 and the remaining ones to 0.

Assigning Items to Components

The optimization procedure could stop here to randomly assign items from pairs to test halves, as is done in the Gulliksen method. However, it is also possible to match the test halves further, for instance, on their average scores or variances. In both cases the problem is a zero-one programming problem again. If the latter option is chosen, the problem is to match the test halves on their sums of the terms $\pi_i(1-\pi_i)\rho_{iX}$ in (3). Since, by definition, there are only two test halves, matching the two sums is equivalent to minimizing the sum with the larger value. Formulating the problem using this minimax criterion has the advantage that it can easily be generalized to other splits than test halves. This generalization will be shown below.

The output of the previous problem is a set of $n/2$ pairs. Let (p,q) denote the p^{th} item ($p = 1, 2$) in the q^{th} pair ($q = 1, \ldots, n/2$) and define a binary decision variable x_{pqr} ($r = 1, 2$) as

$$(8) \qquad x_{pqr} = \begin{cases} 1 & \text{item } (p,q) \text{ is assigned to test half } r, \\ 0 & \text{otherwise.} \end{cases}$$

The assignment problem can be formulated as

(9) minimize z

subject to

$$(10) \qquad \sum_{p=1}^{2} \sum_{q=1}^{n/2} \pi_{pq}(1-\pi_{pq})\rho_{pq}x_{pqr} - z \leq 0 \qquad r = 1, 2$$

$$(11) \qquad \sum_{p=1}^{2} x_{pqr} = 1 \qquad r = 1, 2; \quad q = 1, \ldots, n/2$$

$$(12) \qquad \sum_{r=1}^{2} x_{1qr} = 1 \qquad q = 1, \ldots, n/2$$

$$(13) \qquad x_{pqr} \in \{0, 1\} \qquad p = 1, 2; \quad q = 1, \ldots, n/2$$

where π_{pq} and ρ_{pq} are the item difficulty and discrimination indices for item (p,q). The constraints in (10) ensure that the standard deviations of the two test halves are not larger than the upper bound z minimized in (9). The constraint in (11) requires that the items in a pair are assigned to different test halves each consisting of $n/2$ items, whereas (12) requires that each item is assigned exactly once. The constraints in (11) - (12) could be simplified by replacing (8) by a variable x_{pq} equal to 1 if (p,q)

has to be assigned to the first test half and equal to 0 otherwise, but then the generalization to other splits than tests halves to be presented below is not so obvious.

The same analysis could be done with π_{pq} as coefficients in (10) matching the test halves on their average scores, with weighted combinations $c\pi_{pq} + (1-c)\pi_{pq}(1-\pi_{pq})\rho_{pq}$, $0 \leq c \leq 1$, as coefficients, or with inequality constraints on the averages (variances) added to the model matching the test halves on their variances (averages). All these options are due to the fact that the underlying problem of matching test halves on parallelness is one of multiple-objective decision making. The wealth of choices does not need to bother us much, because the previous pairing of the items already ensures us a high match of the test halves on both their averages and variances before they enter this stage of optimization. In the empirical example below weighted coefficients with $c = .5$ are used. This choice is in the same spirit as the first stage in Gulliksen's method where in (4) π_i and ρ_{ix} are also weighted equally.

Generalization to Other Splits

The above can easily be generalized to other splits than test halves. The case of a split into thirds, for instance, is modeled as follows.

Triples of Items

It is assumed that n is a multiple of three. Then the within-triple "distance" is defined as $\delta_{ijk} = \delta_{ij} + \delta_{ik} + \delta_{jk}$ for all triples (i,j,k), $i \neq j$, $j \neq k$ and $i \neq k$, and the decision variable x_{ijk} is equal to one only if i, j and k are in the same triple and equal to zero otherwise $(i < j < k)$.

The problem is now

$$\text{(14)} \qquad \text{minimize} \quad \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \delta_{ijk} x_{ijk}$$

subject to

$$\text{(15)} \qquad \sum_{j=2}^{k-1} \sum_{i=1}^{j-1} x_{ijk} + \sum_{i=1}^{k-1} \sum_{j=k+1}^{n} x_{ikj} + \sum_{j=k+2}^{n} \sum_{i=k+1}^{j-1} x_{kij} = 1 \qquad k = 1, \ldots, n$$

$$\text{(16)} \qquad x_{ijk} \in \{0, 1\} \qquad i = 1, \ldots, n-2; \; j = i+1, \ldots, n-1; \; k = j+1, \ldots, n$$

where for notational convenience undefined sums in (15) are put equal to zero again. The values in the upper and lower bounds in (5) follow from the requirement that x_{ijk} be defined for $i < j < k$ only.

Assigning Items to Components

If in (9) - (13) the indices run as follows

(17) $p = 1, 2, 3$

$q = 1, \ldots, n/3$

$r = 1, 2, 3$

the model assigns items from triples to test components of size $n/3$.

Conclusion

The above immediately suggests how the model can be generalized to splits into test components of any length.

An Empirical Example

In order to illustrate the procedures, the algorithm by Land and Doig (1960) as implemented in the program LANDO was used together with the item data in Figure 1. The item difficulties and item-test correlations were estimated from a sample of 5418 subjects which is large enough to prevent from capitalizing on chance in the Gulliksen method. The estimates are presented in Table 1.

Insert Table 1 about here

As was clear from the bivariate distribution of the estimates in Figure 1, it is not immediate obvious how all these items should be paired by hand. Table 2 gives the optimal item pairs following (5) - (7). The results of the assignment of the items to test halves according to the

---

Insert Table 2 about here

---

optimization model in (9) - (13), with as coefficients in (10) the equally weighted sum of (2) and (3), are indicated in Table 2 by underscoring the items in the same test half. The results convincingly demonstrate the advantage of optimal assignment over the random assignment that takes place in the original version of the Gulliksen method. For some pairs (e.g., 2-6, 5-15 and 16-17) the within-pair distance is still large in spite of the fact that the pairing was optimal. This implies that there is much space for further optimization. Random assignment makes no systematic use of this but the optimization model in (9) - (13) automatically selects from all possible assignments the one that matches the test halves closest.

Discussion

The idea to estimate lower bounds to the reliability coefficient from a single test administration is not uniquely associated to

Cronbach's coefficient α. It is reminded that other lower bounds to the reliability coefficient exist. One example is Raju's coefficient β_k already referred to earlier. Raju (1982) offers some theory on the theoretical maximum of coefficient α under fixed variance of the test scores. Older examples can be found in Guttman (1945), whereas Bentler and Woodward (1980) (see also ten Berge, Snijders, and Zegers, 1981) derive a whole chain of lower bounds. The idea of maximizing a lower bound is also present in Krammer and van der Linden (1986) who maximize the squared validity coefficient as a lower bound to the reliability coefficient across a set of linear combinations of external variables. All these approaches have different strong and weak points and require more or less intensive computations. It is not the purpose of this paper to replace them by Gulliksen's method. Its main intention is to give this method, which has already gained its place in test theory and practice, a sound computational basis. Besides, the same zero-one programming method can be used in any other situation where classically parallel tests are needed, e.g., for use in pretest-posttest designs in educational research or in testing situations where a secrecy problem exists.

References

Allen, M.J., & Yen, W.M. (1979) Introduction to measurement theory. Monterey, California: Brooks/Cole

Bentler, P.M., & Woodward, J.A. (1980) Inequalities among lower bounds to reliability: With applications to test construction and factor analysis. Psychometrika, 45, 249-268.

Boekkooi-Timminga, E. (1986a) Simultaneous test construction by zero-one programming. Methodika, 1, to appear.

Boekkooi-Timminga, E. (1986b) Algorithms for the construction of parallel tests by zero-one programming. Enschede, The Netherlands: Department of Education, University of Twente.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334

Gulliksen, H. (1950) Theory of mental tests. New York: Wiley.

Guttman, L. (1945) A basis for analyzing test-retest reliability. Psychometrika, 10, 255-282.

Krammer, H.P.M. & van der Linden, W.J. (1986) Reliability estimation using validity coefficients. Journal of Educational Research, 54, 91-94.

Kuder, G.F., & Richardson, M.W. (1937) The theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Land, A.H., & Doig, A. (1960) An automatic method of solving discrete programming problems. Econometrica, 28, 497-520.

Lord, F.M., & Novick, M.R. (1968) Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley.

Raju, N.S. (1977) A generalization of coefficient alpha. Psychometrika, 42, 549-565.

Raju, N.S. (1982) On test homogeneity and maximum KR-20. Educational and Psychological Measurement, 42, 145-152.

ten Berge, J.M.F., Snijders, T.A.B., & Zegers, F.E. (1981) Computational aspects of the greatest lower bound to the reliability and constrained minimum trace factor analysis. Psychometrika, 46, 201-214.

Theunissen, T.J.J.M. (1985) Binary programming and test design. Psychometrika, 50, 411-420.

Theunissen, T.J.J.M., & Verstralen, H.H.F.M. (1985) Algoritmen voor het samenstellen van toetsen Algorithms for test construction . In W.J. van der Linden (Ed.), Moderne methoden voor toetsconstructie en -gebruik. Lisse, The Netherlands: Swets & Zeitlinger.

van der Linden, W.J. & Boekkooi-Timminga, E. (1986) Algoritnmic test design with practical constraints. Enschede, The Netherlands: Department of Education, University of Twente.

Wagner, H.M. (1975) Principles of operations research. London: Prentice/Hall International, Inc.

Table 1

Difficulty and Discrimination Values for the Twenty-item Test

Item	π_i	ρ_{iX}	Item	π_i	ρ_{iX}
1	.85	.39	11	.83	.52
2	.50	.41	12	.68	.54
3	.60	.40	13	.80	.43
4	.66	.60	14	.84	.45
5	.87	.25	15	.86	.34
6	.28	.37	16	.52	.47
7	.87	.40	17	.62	.58
8	.48	.48	18	.61	.40
9	.74	.47	19	.51	.48
10	.65	.60	20	.66	.58

Table 2

Optimal Item Pairs and Test Halves

1 - 7	8 - 19
2 - 6	9 - 13
3 - 18	11 - 14
4 - 10	12 - 20
5 - 15	16 - 17

Note Underscored item numbers in same test half

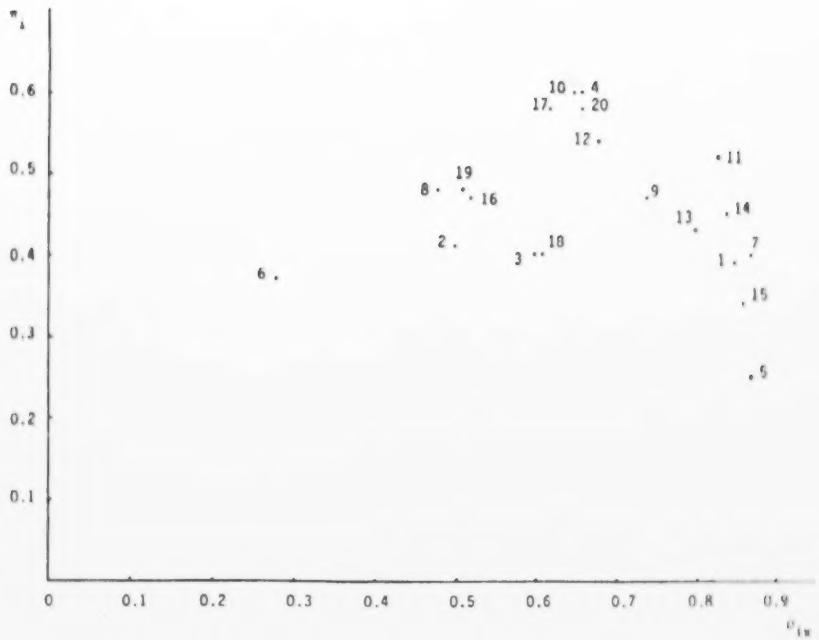

Figure 1. The Gulliksen plot for a twenty-item test.

Authors' Note

This research was supported in part by a grant from the Dutch Organization for Pure Research (ZWO) through the Foundation for Psychological and Psychonomic Research in the Netherlands (PSYCHON). Requests for reprints should be send to the authors address: Faculteit der Toegepaste Onderwijskunde, Universiteit Twente, Postbus 217, 7500 AE Enschede, The Netherlands.

Titles of Recent Research Reports

RR-86-1 W.J. van der Linden, The use of test scores for classification decisions with threshold utility

RR-86-2 H. Kelderman, Item bias detection using the loglinear Rasch model: Observed and unobserved subgroups

RR-86-3 E. Boekkooi-Timminga, Simultaneous test construction by zero-one programming

RR-86-4 W.J. van der Linden, & E. Boekkooi-Timminga, A zero-one programming approach to Gulliksen's random matched subtests method

RR-86-5 E. van der Burg, J. de Leeuw, & R. Verdegaal, Homogeneity analysis with k sets of variables: An altenating least squares method with optimal scaling features

RR-86-6 W.J. van der Linden, & T.J.H.M. Eggen, An empirical Bayes approach to item banking

RR-86-7 E. Boekkooi-Timminga, Algorithms for the construction of parallel tests by zero-one programming

RR-86-8 T.J.H.M. Eggen, & W.J. van der Linden, The use of models for paired comparisons with ties

Research Reports can be obtained at costs from Mediatheek, Faculteit Toegepaste Onderwijskunde, Universiteit Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.

A publication by
the Department of Education
of the University of Twente,
P.O. Box 217,
7500 AE Enschede,
The Netherlands

Department of Education